

**Global Assessment of Functioning (GAF) Scores:
Clinicians' Ratings do not Predict Substance Abuse Patients'
1-Year Treatment Outcomes**

Rudolf H. Moos, Lucie McCoy, and Bernice S. Moos

Center for Health Care Evaluation
and Program Evaluation and Resource Center
Veterans Affairs Health Care System
Palo Alto, California

**A Report From the Substance Abuse Module,
Quality Enhancement Research Initiative (QUERI)**

Contents

Executive Summary.....	i
Introduction.....	1
Correlates and Predictors of GAF Ratings.....	1
GAF Ratings as Predictors of Treatment Outcomes.....	2
Specific Questions Addressed.....	3
Method.....	3
Patients.....	4
Measures.....	4
Analysis Plan.....	5
Results.....	6
Determinants of Clinicians' GAF Ratings of Patients' Current Functioning.....	7
Clinicians' Ratings of Patients' Current Functioning and Treatment Outcome.....	9
Clinicians' Ratings of Patients' Highest Level of Functioning.....	10
Comment.....	11
Determinants of GAF Ratings.....	11
GAF Ratings as Predictors of Outcome.....	11
Conclusions and Recommendations.....	12
Acknowledgments.....	14
References.....	15

Titles of Tables

Table 1.	Demographic and Diagnostic Characteristics of Patients with Different Levels of Global Impairment	6
Table 2.	Psychological, Social, and Occupational Functioning at Treatment Intake of Patients with Different Levels of Global Impairment.....	7
Table 3.	Multiple Regression Analyses to Identify Independent Predictors of GAF Ratings of Patients' Current and Highest Level of Functioning	8
Table 4.	Current Global Assessment of Functioning (GAF) Ratings as Predictors of Patients' 1- year Outcomes.....	9

Executive Summary

Background. The Global Assessment of Functioning (GAF) Scale is a standard part of the American Psychiatric Association's (APA) multidimensional diagnostic system. VA policy requires clinicians to use the GAF, which assesses mental health patients' current level of functioning on a 1-100 scale, as part of the standard diagnostic procedure. GAF scores are recorded in the Veterans Health Information System and Technology Architecture (VistA) and in the nationwide VA healthcare utilization database.

Objective. To assess whether readily available information about patients' current level of psychosocial functioning, as reflected in clinicians' GAF ratings, can be used as a predictor of substance abuse patients' treatment outcomes.

Method. We identified a sample of 1,688 VA patients with substance use disorders, many of whom also had psychiatric disorders; examined the determinants of GAF ratings; and focused on how well these ratings predicted patients' 1-year symptom and psychosocial functioning outcomes.

Results. Patients' clinical diagnoses and psychiatric symptoms were stronger predictors of GAF ratings than was their current level of social and occupational functioning. Moreover, GAF ratings were only minimally associated with patients' 1-year psychological, social, and occupational functioning outcomes.

Conclusions. These findings raise serious questions about the conceptual and clinical value of the standard method of assessing psychiatric and substance abuse patients' global level of functioning. They imply that, as currently employed in VA, clinicians' GAF ratings of substance abuse patients' global functioning cannot be used as adequate predictors of patients' treatment outcomes.

Recommendations. Work with the American Psychiatric Association and other professional mental health organizations to identify a better standard measure to assess patients' psychosocial functioning; train VA clinicians in the use of the chosen measure and conduct an evaluation of its reliability and outcome-based validity.

Introduction

The Global Assessment of Functioning (GAF) Scale, which obtains information for Axis V of DSM-IV, is the standard method used to assess the clinician's judgment of a patient's overall level of functioning (APA, 1994). As such, the GAF probably is the single most widely used rating scale to assess impairment among patients with psychiatric and/or substance use disorders (Piersma & Boes, 1997).

VA policy requires clinicians to use the GAF, which assesses mental health patients' current level of functioning on a 1-100 scale, as part of the standard diagnostic procedure. GAF scores are recorded in the Veterans Health Information System and Technology Architecture (VistA) and in the nationwide VA healthcare utilization database (VHA Directive 97-059).

The GAF requires the clinician to make an overall judgment about a patient's current and highest (for several months during the past year) level of psychological, social, and occupational functioning. These ratings are based on a 1-10 (severely impaired) to 81-90 (superior functioning) scale in DSM-III-R and on a 1-100 scale in DSM-IV. The GAF is a revised version of the Global Assessment Scale (GAS) that was used as Axis V of DSM-III to assess a patient's overall functioning (APA, 1987; Endicott et al., 1976). In turn, the GAS began as a slightly revised version of the original 100-point Health-Sickness Rating Scale (Luborsky, 1962).

According to the DSM-IV manual, the clinician's ratings of current functioning reflect the patient's need for treatment and should be used for treatment planning (APA, 1994). Ratings of the patient's current and highest level of functioning are expected to be of value in predicting treatment outcome (Bodland et al., 1994; Phelan, Wykes, & Goldman, 1994).

Surprisingly, however, there is relatively little empirical information on the adequacy of the GAF for these purposes. Moreover, although the GAF appears to be reasonably reliable and valid in a research context (e.g., Tracey et al., 1997), there is virtually no data on the value of the most prevalent use of the GAF, which is as a standard part of experienced clinicians' regular diagnostic assessments. A recurrent conceptual problem is how well a unidimensional rating scale, such as the GAF, reflects or predicts patients' combined symptom and social and occupational functioning outcomes.

Correlates and Predictors of GAF Ratings

One important issue involves the determinants of GAF scores; that is, the main factors clinicians consider when they rate patients' current global functioning. In general, GAF scores are relatively independent of sociodemographic factors (Coffey, Jones, & Thornicroft, 1996; Roy-Byrne et al., 1996; Skodol et al., 1988a, 1988b), suggesting that clinicians' ratings of functioning do not rely on patients' age, gender, marital status, and the like.

Several studies have identified predictable concurrent associations between the severity of patients' symptoms and clinicians' ratings of global impairment (Roy-Byrne et al., 1996). For example, Endicott et al. (1976) found that patients' symptoms of cognitive disorganization, hallucinations, delusions, suspiciousness, and inappropriate appearance were associated with more global impairment.

Lower GAF ratings also have been linked with clinician- and self-rated depression, suicidal ideation, and lack of self-esteem (Coulehan et al., 1997; Hall, 1995; Mueser et al., 1997; Van Gastel, Schotte, & Mae, 1997), more cognitive impairment (Meltzer et al., 1996), and higher self-rated severity of illness (Hall, 1995). With respect to diagnoses, patients with Axis I disorders, especially psychoses, tend to be rated as more impaired (i.e., to obtain lower GAF scores) than do patients with only Axis II disorders (Coffey et al., 1996; Phelan et al., 1994; Skodol et al., 1988b).

A few studies have found that GAF ratings are predictably correlated with indices of social functioning, such as limited social networks and need for support (Jones et al., 1995; Phelan et al., 1994; Westermeyer & Neider, 1988), and with work adjustment (Roy-Byrne et al., 1996). In general, however, GAF ratings seem to be more closely associated with patients' symptoms than with their social and occupational functioning (Brekke, 1992; Skodol et al., 1988b). Roy-Byrne et al. (1996) found that a GAF modified to focus specifically on social and occupational functioning, and not on clinical symptoms, still was more closely related to patients' clinical symptoms than to these other indices of functioning.

GAF Ratings as Predictors of Treatment Outcomes

A number of studies have shown that clinicians' ratings of patients' global functioning index improvements during and following treatment (Furukawa et al., 1995; Hawthorne et al., 1994; Howes et al., 1997; Kocsis et al., 1997; Linehan et al., 1994; Piersma & Boes, 1997; Rund et al., 1994; Walton et al., 1996). In most of these studies, however, GAF ratings were completed by treatment providers themselves, raising the question of potential bias.

More important, we found only two studies that examined the association between GAF ratings during treatment and independent information about posttreatment outcomes. Higher GAF scores were moderate predictors of more employment, more hours worked, and more earned income (Mueser et al., 1997; Vetter & Keller, 1996). In an earlier study, Mellsop, Peace, & Fernando (1987) showed an association between GAF-rated impairment at treatment intake and poorer 6-month outcomes; however, this finding held only for schizophrenic patients (see also Beiser et al., 1988).

Specific Questions Addressed

We focus here on a nationwide sample of VA patients and address two related questions:

(1) What are the determinants of clinicians' ratings of patients' current global functioning? Specifically, are patients' diagnostic characteristics and clinical symptoms the most important predictors of GAF scores? To what extent do patients' social and occupational functioning at intake to treatment predict their GAF scores?

(2) How well does the GAF predict patients' 1-year symptom and social and occupational functioning outcomes? Do GAF ratings add predictive value over and above information about diagnostic factors that is readily available in administrative databases? Do they add predictive value over and above the baseline value of the outcome criterion? We also examine these issues with respect to clinicians' ratings of patients' highest level of functioning in the past year.

Method

The data were drawn from a nationwide study of the 1-year outcome of treatment for patients with substance use disorders, many of whom also had concomitant psychiatric disorders. In the overall study, more than 3,600 VA patients completed an Intake Information Form (IIF) at the time they entered treatment in one of a representative sample of 15 VA substance abuse programs. The IIF assessed sociodemographic characteristics, indices of substance use, and patients' psychological, social, and occupational functioning (to be described below).

An attempt was made to contact each patient about 12 months after discharge from the index episode to complete a Follow-up Information Form (FIF), which assessed identical content areas as the IIF. A total of 86 participants died during the follow-up period. Of the remaining patients, 3,018 (84%) completed the follow-up, which was administered by mail and in-person and telephone interviews. (For more details about the patients and assessment methods, see Moos et al., 1999; Ouimette, Finney, & Moos, 1997.)

There were no differences between patients who were and those who were not successfully followed on demographic factors (age, education, ethnic background, income, employment status) or alcohol dependence symptoms at intake to treatment. We used biological tests (e.g., urine/blood/breath samples) to verify a subset of patients' self-reports of substance use. Patients' reports of abstinence from alcohol and drugs were significantly associated with negative alcohol (95%) and drug (86%) tests (Ouimette et al., 1997).

Patients

We focus here on the subset of 1,688 followed patients for whom the DSM-III-R version of the GAF was completed during the inpatient treatment episode. On average, these patients were 43 years old and had completed almost 13 years of education. A total of 51% of the patients were Caucasian, 43% were African-American, 3% were Hispanic/Latino, and the remaining 2% were Asian, Native American, or other. Only 20% of the patients were married. These demographic characteristics were comparable to those of the larger group of followed patients from which this subsample was drawn.

With respect to their substance use disorder diagnoses, 37% of the patients were alcohol dependent only, 12% were drug dependent only, and 51% were dependent on both alcohol and drugs. In addition, 41% of the patients had a concomitant psychiatric diagnosis; 13% had an Axis I psychiatric diagnosis (5% had a psychosis and 7% had posttraumatic stress disorder) and 24% had a personality disorder.

Measures

Patients' baseline functioning and 1-year outcomes were assessed with respect to eight indices: Three indices of psychological functioning focused on:

- emotional distress as assessed by responses on 5-point scales varying from "not at all" to "extremely" to 12 items (such as "feelings of worthlessness", "thoughts of ending your life", and "spells of terror or panic") drawn from the Depression and Anxiety Scales of the Brief Symptom Inventory (BSI; Derogatis, 1993);
- psychiatric symptoms as assessed by responses on comparable 5-point scales to 10 items (such as "feeling that you are watched or talked about by others", and "the idea that someone else can control your thoughts") drawn from the BSI Paranoid Ideation and Psychoticism Scales; and
- substance use problems, as assessed by 18 items rated on 4-point scales varying from "never" to "frequently" reflecting a standard set of problems due specifically to alcohol and/or drug use, such as "legal problems", "problems with your job", and "problems with your neighbors."

Three indices of social functioning focused on:

- residential stability as assessed by living in a stable setting in the community for most of the past 12 months and never or seldom losing a place to live in the three months prior to follow-up (yes/no);
- the number of friends with whom the patient felt at ease; and
- the quality of relationships with friends as measured by 5-point ratings on six items

drawn from the Life Stressors and Social Resources Inventory (Moos & Moos, 1994) such as “Do you confide in any of your friends?” and “Do your friends really understand how you feel about things?”

In addition, we obtained information on two indices of occupational functioning: (1) part-time or full-time employment status (yes/no), and (2) annual income.

Analysis Plan

Although GAF ratings are typically made on a continuous 100-point scale, the rating criteria are described in 9 broad categories (APA, 1994) and many researchers have combined GAF scores into a smaller number of levels of functioning (Hall, 1995; Mezzich, Fabrega, & Coffman, 1987; Rabinowitz et al., 1994; Schrader, Gordon, & Harcourt, 1986; Thompson et al., 1992; Van Gastel et al., 1997). Accordingly, we divided patients into five groups on the basis of clinicians’ ratings of their current functioning: (1) pervasive impairment (GAF scores of 1-40), (2) serious impairment (GAF scores of 41-50), (3) moderate impairment (GAF scores of 51-60), (4) mild impairment (GAF scores of 61-70), and (5) minimal impairment (GAF scores of 71-90).

To examine the extent to which baseline demographic and diagnostic variables and indices of psychological, social, and occupational functioning differentiated the five groups, we conducted analyses of variance (ANOVAs) for continuous variables and chi square for dichotomous variables.

We then conducted multiple regression analyses to identify the best independent predictors of clinicians’ GAF ratings of patients’ current global functioning, as reflected in the five GAF groups.

Next, we calculated correlations between the GAF ratings and 1-year outcomes; subsequent analyses controlled for patients’ diagnoses and prior treatment and the baseline value of the outcome criterion. We repeated these analyses to examine how well clinicians’ ratings of patients’ highest levels of functioning in the past year predicted patients’ treatment outcomes.

Results

The five GAF groups did not differ significantly in demographic characteristics, but they varied substantially in diagnoses. Specifically, more impaired patients were more likely to have both alcohol and drug diagnoses and Axis I psychiatric diagnoses (Table 1). More impaired patients were also more likely to have medical diagnoses (not shown). In addition, they were more likely to have had inpatient mental health care in the prior year.

Table 1. Demographic and Diagnostic Characteristics of Patients with Different Levels of Global Impairment

Demographic Factors	Level of Global Impairment					F or Chi Square
	Pervasive (N=105)	Serious (N=355)	Moderate (N=589)	Mild (N=521)	Minimal (N=118)	
Age (years)	42.9	43.7	43.3	42.3	43.1	1.33
Married (%)	14.3	22.2	18.3	18.4	18.5	3.49
Education (years)	12.5	12.7	12.6	12.7	12.4	<1
Caucasian (%)	52.4	54.9	53.0	46.8	44.1	9.06
<u>Dx and Prior Rx (%)</u>						
Both Alcohol and Drug Dx	64.8	55.1	51.3	48.2	37.3	20.88**
Psychiatric Dx	47.6	42.3	47.8	30.1	39.8	38.84**
Axis I Dx	24.8	14.9	14.8	10.4	5.9	23.08**
Psychoses	10.5	4.2	6.1	3.6	0.8	15.24**
PTSD	13.3	8.2	7.8	5.8	4.2	9.68**
Inpatient Episode in Prior Year (%)	31.4	34.4	36.1	26.1	25.4	16.02**

**p < .01

Determinants of Clinicians' GAF Ratings of Patients' Current Functioning

Consistent with the diagnostic differences among the five GAF groups, more impaired patients had higher scores on emotional distress and psychiatric symptoms, and had more substance use problems. The GAF groups differed in social and occupational functioning: more impaired patients were less stable residentially, had fewer friends, were less likely to be employed, and had lower incomes.

Table 2. Psychological, Social, and Occupational Functioning at Treatment Intake of Patients with Different Levels of Global Impairment

Area of Functioning	Level of Global Impairment					F or Chi Square
	Pervasive (N=105)	Serious (N=355)	Moderate (N=589)	Mild (N=521)	Minimal (N=118)	
<u>Psychological</u>						
Distress (0-48)	22.8	21.6	20.9	18.2	15.1	13.60**
Psychiatric Symptoms (0-40)	16.5	15.1	14.8	12.9	11.3	9.67**
Substance Use Problems (0-72)	27.3	25.6	23.9	21.1	19.2	11.81**
<u>Social</u>						
Residential Stability (% yes)	51.4	62.8	61.5	68.7	75.4	5.24**
Number of Friends	1.6	2.1	2.0	2.1	2.2	2.52*
Friendships (0-24)	12.2	13.4	13.1	13.2	14.1	1.75
<u>Occupational</u>						
Employed (%)	20.0	22.0	19.0	27.5	32.2	17.32**
Income (in \$1,000s)	4.7	6.3	6.4	7.6	8.4	4.07**

*p < .05; **p < .01

We conducted multiple regression analysis to identify the best independent predictors of GAF ratings. We entered the diagnostic and functioning indices that were significantly associated with GAF ratings, as shown in Tables 1 and 2. Patients' substance use, Axis I psychiatric, and medical diagnoses each significantly predicted more GAF impairment, as did patients' emotional distress and substance use problems (Table 3). Taken together, however, these variables accounted for only 8% of the variance in clinicians' GAF ratings of patients' current global functioning.

Table 3. Multiple Regression Analyses to Identify Independent Predictors of GAF Ratings of Patients' Current and Highest Level of Functioning

Predictors	Current Global Functioning (N = 1,688)	Highest Level of Functioning (N = 1,282)
<u>Diagnoses</u>		
Both Alcohol and Drug	-.09**	-.02
Axis I Psychiatric	-.08**	-.06*
Medical	-.15**	-.10**
<u>Psychological Functioning at Intake</u>		
Distress	-.08**	-.10**
Substance Use Problems	-.10**	-.05
<u>Occupational Functioning at Intake</u>		
Employed	.02	.07*
Annual Income	.07*	.01
Multiple R	.28**	.21**
R ²	.08	.04

*p < .05;

**p < .01

Clinicians' Ratings of Patients' Current Functioning and Treatment Outcome

Next, we examined whether GAF ratings obtained during treatment were associated with patients' 1-year outcomes. As shown by the correlations in the first column in Table 4, patients who were rated as more impaired on the GAF reported more distress and psychiatric symptoms and more substance use problems at 1 year. These patients also were less likely to be in stable residential situations. However, GAF ratings were not related to patients' 1-year friendships, employment, or income. Moreover, none of the correlations accounted for more than about 1% of the variance in patients' outcomes

Table 4. Current Global Assessment of Functioning (GAF) Ratings as Predictors of Patients' 1-Year Outcomes

1-Year Outcome	GAF Rating of Current Functioning	GAF Rating Controlling Dx and Prior Rx	GAF Rating Controlling Baseline Value of Outcome
<u>Psychological Functioning</u>			
Distress	-.09**	-.06**	.00
Psychiatric Symptoms	-.11**	-.08**	-.02
Substance Use Problems	-.08**	-.05*	.00
<u>Social Functioning</u>			
Residential Stability	.07**	.04	.02
Number of Friends	.03	.02	.03
Friendships	.02	.01	.01
<u>Occupational Functioning</u>			
Employed	.04	.01	-.01
Income	.04	.02	-.01

*p < .05; **p < .01

These findings show some very moderate associations between GAF ratings of patients' current functioning and their treatment outcomes 1 year later. We obtained essentially identical findings when we used continuous GAF scores rather than categorizing the values into five groups. We also conducted two sets of multiple regression analyses to find out whether GAF ratings added predictive information over and above (1) readily available information about patients' diagnoses and prior treatment, and (2) the value at intake of the 1-year outcome criterion.

After controlling for patients' diagnoses and prior treatment, GAF ratings still added a small amount of predictable variance to 1-year distress, psychiatric symptoms, and substance use problems (partial correlations in the second column in Table 4). As expected, each of the 8 outcome criteria was predicted significantly by its value at intake (r s ranged from .27 to .50; all p s < .01).

After controlling for the baseline value of the outcome criterion, GAF ratings were not significantly associated with any of the 1-year outcomes (partial correlations in the third column in Table 4). Moreover, each of the eight intake values predicted the equivalent outcome criterion better than the GAF rating did. These findings held in separate analyses on the subset of patients who had psychiatric diagnoses.

We also examined the intercorrelations among the eight outcome criteria. The indices of psychological and social functioning were only moderately intercorrelated (average $r = .34$). The intercorrelations among the indices of psychological and occupational functioning (average $r = .24$) and those among the indices of social and occupational functioning (average $r = .15$) were even lower. Accordingly, a single rating, such as the GAF rating, cannot adequately predict these relatively independent outcome criteria.

Clinicians' Ratings of Patients' Highest Level of Functioning

Clinicians' also rated most patients' ($N = 1,282$) highest level of functioning in the past year. Ratings of patients' current and highest levels of functioning were substantially correlated ($r = .72$; $p < .01$). Patients' diagnoses and symptoms were more strongly associated with their highest level of functioning than were indices of social and occupational adaptation (Table 3). With respect to the predictive value of patients' highest level of functioning, these ratings were significantly associated with only the three symptom outcomes (r s = $-.08$, $-.10$, and $-.07$ for emotional distress, psychiatric symptoms, and substance use problems, respectively; all P s < .05). After controlling for the baseline value of the outcome criterion, none of these relationships was still significant.

Comment

Our findings raise serious questions about the conceptual and empirical basis of experienced clinicians' routine ratings of patients' current and highest levels of global functioning for Axis V of DSM-IV.

Determinants of GAF Ratings

Clinicians' ratings of patients' current levels of global functioning appear to be based primarily on patients' diagnoses and psychiatric symptoms. Specifically, patients with both alcohol and drug diagnoses, Axis I psychiatric diagnoses, and medical diagnoses are likely to be rated as more impaired, as are patients who report more emotional distress, psychiatric symptoms, and substance use problems. Once these clinical factors have been considered, indices of social and occupational functioning add little if any predictable variance to GAF ratings.

These findings hold for patients' highest levels of functioning, as well as for their current levels. The results support prior work on the associations between patients' symptoms and Axis V global impairment ratings (Coulehan et al., 1997; Endicott et al., 1976; Hall, 1995; Muesser et al., 1997; Van Gastel et al., 1997), and imply that such ratings contain little if any information about patients' social or occupational functioning that is independent of clinicians' judgments about the severity of patients' symptoms (Brekke, 1992; Roy-Byrne et al., 1996; Skodol et al., 1988b).

Consistent with several other studies (Dufton & Siddique, 1992; Piersma & Boes, 1995; Sullivan & Grubea, 1991), we found only minimal associations between patients' self-rated symptoms and clinicians' ratings of patients' current and highest level of functioning. This finding raises further questions about the determinants and value of GAF ratings, especially given that patients' self-rated symptoms at baseline predict 1-year symptom and functioning outcomes much better than do clinicians' ratings of patients' global impairment.

GAF Ratings as Predictors of Outcome

We found little if any relationship between ratings of patients' current or highest level of global functioning and psychological, social, or occupational functioning at a 1-year follow-up. The few significant relationships we identified accounted for less than 1% of the variance in the 1-year outcomes and became nonsignificant after controlling for the baseline value of the relevant outcome criterion. Moreover, as might be expected, these baseline values were substantially better predictors of 1-year outcomes than were the GAF ratings.

Because patients' usually return to their prior level of functioning after an episode of acute illness, ratings of patients' highest level of functioning during the prior year are thought to have some prognostic value. However, we found essentially no relationship

between patients' rated highest level of functioning and their 1-year outcomes. In conjunction with the lack of prior positive findings linking GAF ratings to outcomes (Goldman, Skodol, & Lave, 1992), these findings cast serious doubt on the empirical value of global impairment ratings for predicting treatment outcome.

We also found that global impairment ratings did not predict patients' likelihood of readmission or length of readmission. Although lower GAF scores were associated with longer hospital episodes (see also Dufton & Siddique, 1992; Gordon & Gordon, 1987; Rabinowitz et al., 1994; Tucker et al., 1987; Vetter & Koller, 1996), there was no evidence for a link between more clinician-rated impairment and readmission or length of readmission. These findings probably reflect the complexity of the determinants of readmission. They imply that the GAF does not fulfill one of its primary aims, which is to predict patients' future use of inpatient treatment services.

Conclusions and Recommendations

It is important to recognize the limitations of this project. The data reflect the use of the GAF in a nationwide system of care with many different clinicians who varied in their level of experience and expertise. Although the GAF ratings were completed as part of a standard clinical diagnostic interview, they were not obtained under controlled conditions with specific training to enhance reliability. Another limitation is that the data were obtained only from men and from patients with substance use disorders, and were drawn only from the VA system of care.

Further research may be warranted to find out whether a well-structured program of training can enhance the reliability and validity of the GAF in the VA and in other clinical settings. Fundamentally, however, our findings are comparable to those of several prior studies, and seem to be representative of the characteristics and value of the GAF in the actual clinical situation.

More than 25 years of experience have shown only limited value of a unidimensional rating of global functioning based on psychological, social, and occupational criteria. Such unidimensional ratings represent a fundamentally conceptually flawed procedure, because these three areas of functioning at best are only moderately interrelated. We identified only moderate interrelationships among these indices of functioning here, as well as in earlier studies of substance abuse patients (Moos, Finney, & Cronkite, 1990). Similar findings have been obtained among schizophrenic patients, diverse sets of psychiatric patients, and untreated individuals in the community (Dohrenwend et al., 1983; Strauss & Carpenter, 1977). In addition, symptoms tend to change more quickly during treatment than does functioning (Gordon et al., 1988).

Conclusions. We conclude that the time has come to curtail the use of a single global impairment rating for Axis V of DSM-IV; to reconceptualize Axis V as tri-dimensional with separate indices of psychological, social, and occupational functioning; and to identify simple, behaviorally based indicators of patients' adaptation in each of these three areas.

For example, psychological functioning could be assessed by a brief index drawn from the BSI (Derogatis, 1993), social functioning by residential stability and/or the quality of relationships with friends (Moos & Moos, 1994), and occupational functioning by whether or not the patient is employed or the number of weeks of employment in the past three months.

Initial findings based on a potential modification of the GAF, the Social and Occupational Functioning Assessment Scale (SOFAS) (APA, 1994; Goldman et al., 1992), indicate that it still primarily reflects symptoms rather than other domains of adaptation (Roy-Byrne et al., 1996, but see Patterson & Lee, 1995). This may be due to the fact that the SOFAS is based closely on the GAF, does not include behaviorally based indicators of functioning to guide the rater, combines the essentially independent domains of social and occupational functioning, and requires raters to consider only impairments due to mental and physical health problems.

Recommendations. VA program managers and providers should work with the American Psychiatric Association and other professional mental health organizations to identify a better standard measure to assess patients' psychosocial functioning. After a preferred measure is chosen, the VA needs to train clinicians in the use of the measure and to conduct an evaluation of its reliability and outcome-based validity.

The growing emphasis on accountability and outcomes monitoring in mental health services highlights the importance of developing reliable and valid measures to reflect the multidimensionality of patients' functioning in psychosocial, social, and occupational domains. A common set of definitions and criteria for Axis I and II diagnoses has led to substantial advances in characterizing patients' disorders and symptoms. We need to fulfill the promise of a true multi-axial assessment system, and establish conceptually and empirically independent measures of patients' impairments and adaptation in different functional domains.

Acknowledgments

This report was prepared for the Substance Abuse Module of the VA Quality Enhancement Research Initiative (QUERI). Preparation of the report was supported by the Department of Veterans Affairs Office of Research and Development (Health Services Research and Development Service) and Mental Health Strategic Health Group. John Finney, Keith Humphreys, Paige Crosby Ouimette, and Kathleen Schutte made helpful comments on an earlier draft of this report. An adapted version of the report is in press in the Journal of Clinical Psychology.

References

- American Psychiatric Association (1987). Diagnostic and statistical manual of mental disorders (3rd ed.). Washington, DC: Author.
- American Psychiatric Association (1994). Diagnostic and statistical manual of mental disorders (4th ed.). Washington, DC: Author.
- Beiser, M., Fleming, J.A.E., Fleming, M.B., Lacono, W.G., & Lin, T (1988). Refining the diagnosis of schizophreniform disorder. American Journal of Psychiatry, 145, 695-700.
- Bodland, O., Kullgren, G., Ekselius, L., Lindstrom, E., & Von Knorring, L. (1994). Axis V - Global Assessment of Functioning Scale: Evaluation of a self-report version. Acta Psychiatrica Scandinavica, 90, 342-347.
- Brekke, J. S. (1992). An examination of the relationships among three outcome scales in schizophrenia. Journal of Nervous Mental Disease, 180, 162-167.
- Coffey, M., Jones, S., & Thornicroft, G. (1996). A brief mental health outcome scale: relationships between scale scores and diagnostic/sociodemographic variables in the long-term mentally ill. Journal of Psychiatric Mental Health Nursing, 3, 89-93.
- Coulehan, J., Schulberg, H., Block, M., Madonia, M., & Rodriguez, E. (1997). Treating depressed primary care patients improves their physical, mental, and social functioning. Archives of Internal Medicine, 157, 1113-1120.
- Derogatis, L. (1993). Brief Symptom Inventory: Administration, scoring, and procedures manual (3rd ed.). Minneapolis, MN: National Computer Systems.
- Dohrenwend, B.S., Dohrenwend, B.P., Link, B., & Levav, I. (1983). Social functioning of psychiatric patients in contrast with community cases in the general population. Archives of General Psychiatry, 40, 1174-1182.
- Dufton, B. D., & Siddique, C. M. (1992). Measures in the day hospital. I. The Global Assessment of Functioning Scale. International Journal of Partial Hospitalization, 8, 41- 49.
- Endicott, J., Spitzer, R., Fleiss, J., & Cohen, J. (1976). The Global Assessment Scale: A procedure for measuring overall severity of psychiatric disturbance. Archives of General Psychiatry, 33, 766-771.
- Furukawa, T., Awaji, R., Nakazato, H., & Sumita, Y. (1995). Predictive validity of subtypes of chronic affective disorders derived by cluster analysis. Acta Psychiatrica Scandinavica, 91, 379-385.

- Goldman, H. H., Skodol, A. E., & Lave, T. R. (1992). Revising axis V for DSM-IV: A review of measures of social functioning. American Journal of Psychiatry, 149, 1148-1156.
- Gordon, R. E., & Gordon, K. K. (1987). Relating axes IV and V of DSM-III to clinical severity of psychiatric disorders. Canadian Journal of Psychiatry, 32, 423-424.
- Gordon, R. E., Plutzky, M., Gordon, K. K., & Guerra, M. (1988). Using the Axis V Scale to evaluate therapeutic outcome of psychiatric treatment. Canadian Journal of Psychiatry, 33, 194-196.
- Hall R. C. W. (1995). Global Assessment of Functioning: A modified scale. Psychosomatics, 36, 267-275.
- Hawthorne, W. B., Fals-Stewart, W., & Lohr, J. B. (1994). A treatment outcome study of community-based residential care. Hospital Community Psychiatry, 45, 152-155.
- Howes, J. L., Haworth, H., Reynolds, P., & Kavanaugh, M. (1997). Outcome evaluation of a short-term mental health day treatment program. Canadian Journal of Psychiatry, 42, 502-508.
- Jones, S. H., Thornicroft, G., Coffey, M., & Dunn, G. (1995). A brief mental health outcome scale: Reliability and validity of the Global Assessment of Functioning (GAF). British Journal of Psychiatry, 166, 654-659.
- Kocsis, J. H., Zisook, S., Davidson, J., Shelton, R., Yonkers, K., Hellerstein, D. J., Rosenbaum, J., & Halbreich, U. (1997). Double-blind comparison of sertraline, imipramine, and placebo in the treatment of dysthymia: Psychosocial outcomes. American Journal of Psychiatry, 154, 390-395.
- Linehan, M. M., Tutek, D.A., Heard, H.L., & Armstrong, H.E. (1994). Interpersonal outcome of cognitive behavioral treatment for chronically suicidal borderline patients. American Journal of Psychiatry, 151, 1771-1776.
- Luborsky, L. (1962). Clinicians' judgments of mental health: A proposed scale. Archives of General Psychiatry, 7, 407-417.
- Mellsop, G., Peace, K., & Fernando, T. (1987). Pre-admission adaptive functioning as a measure of prognosis in psychiatric inpatients. Australian and New Zealand Journal of Psychiatry, 21, 539-544.
- Meltzer, H.Y., Thompson, P.A., Myung, A.L., & Ranjan, R. (1996). Neuropsychologic deficits in schizophrenia: Relation to social function and effect of antipsychotic drug treatment. Neuropsychopharmacology, 14, 27S-33S.

- Mezzich, J.E., Fabrega, H., & Coffman, G.A. (1987). Multiaxial characterization of depressive patients. Journal of Nervous and Mental Disease, 175, 339-346.
- Moos R., Finney, J., & Cronkite, R. (1990). Alcoholism treatment: Context, process, and outcome. New York: Oxford.
- Moos, R., Finney, J., Ouimette, P., & Suchinsky, R. (1999). A comparative evaluation of substance abuse treatment: 1. Treatment orientation, amount of care, and 1-year outcomes. Alcoholism: Clinical and Experimental Research, 23, 529-536.
- Moos, R., & Moos, B. (1994). Life Stressors and Social Resources Inventory: Adult form manual. Odessa, FL: Psychological Assessment Resources, Inc.
- Mueser, K. T., Becker, D.R., Torrey, W.C., Xie, H., Bond, G.R., Drake, R.E., & Dain, B. J. (1997). Work and nonvocational domains of functioning in persons with severe mental illness: A longitudinal analysis. Journal of Nervous and Mental Disease, 185, 419-426.
- Ouimette, P. C., Finney, J. W., & Moos, R. (1997). Twelve step and cognitive behavioral treatment for substance abuse: A comparison of treatment effectiveness. Journal of Consulting Clinical Psychology, 65, 220-234.
- Patterson, D. A., & Lee. M. (1995). Field trial of the Global Assessment of Functioning Scale - Modified. American Journal of Psychiatry, 152, 1386-1388.
- Phelan, M., Wykes, T., & Goldman, H. (1994). Global function scales. Social Psychiatry and Psychiatric Epidemiology, 29, 205-211.
- Piersma, H. L., & Boes, J. L. (1995). Agreement between patient self-report and clinician rating: Concurrence between the BSI and the GAF among psychiatric inpatients. Journal of Clinical Psychology, 51, 153-157.
- Piersma, H. L., & Boes, J. L. (1997). The GAF and psychiatric outcome: A descriptive report. Community Mental Health Journal, 33, 35-40.
- Rabinowitz, J., Modai, I., & Inbar-Saban, N. (1994). Understanding who improves after psychiatric hospitalization. Acta Psychiatrica Scandinavia, 89, 152-158.
- Roy-Byrne P., Dagadakis C., Unutzer J., & Ries R. (1996). Evidence for limited validity of the revised Global Assessment of Functioning Scale. Psychiatric Services, 47, 864-866.

- Rund, B. R., Sollien, T., Fjell, A., Borchgrevink, T., Hallert, M., & Naess, P. O. (1994). The Psychosis Project: outcome and cost-effectiveness of a psychoeducational treatment programme for schizophrenic adolescents. Acta Psychiatrica Scandinavica, 89, 211-218.
- Schrader, G., Gordon, M., & Harcourt, R. (1986). The usefulness of DSM-III Axis IV and Axis V assessments. American Journal of Psychiatry, 143, 904-907.
- Skodol, A. E., Link, B. G., Shrout, P. E., & Horwath, E. (1988a). The revision of axis V in DSM-III-R: Should symptoms have been included? American Journal of Psychiatry, 145, 825-829.
- Skodol, A. E., Link, B. G., Shrout, P. E., & Horwath, E. (1988b). Toward construct validity for DSM-III Axis V. Psychiatry Research, 24, 13-23.
- Strauss, J. S., & Carpenter, W. T. (1977). Prediction of outcome in schizophrenia. Archives of General Psychiatry, 34, 159-163.
- Sullivan, C. W., & Grubea, J. M. (1991). Who does well in a day treatment program? Following patients through 6 months of treatment. International Journal of Partial Hospitalization, 7, 101-10.
- Thompson, J.W., Burns, B.J., Goldman, H.H., & Smith, J. (1992). Initial level of care and clinical status in a managed mental health program. Hospital and Community Psychiatry, 43, 599-603.
- Tracey, K., Adler, L. A., Rotrosen, J., Edson, R., & Lavori, P. Interrater reliability issues in multicenter trials, part 1: Theoretical concepts and operational procedures used in Department of Veterans Affairs cooperative study, #394. Psychopharmacology Bulletin, 33, 53-57.
- Tucker, L., Bauer, S. F., Wagner, S., Harlam, D., & Sher, I. (1987). Long-term hospital treatment of borderline patients: A descriptive outcome study. American Journal of Psychiatry, 144, 1443-1448.
- Van Gastel, A., Schotte, C., & Maes, M. (1997). The prediction of suicidal intent in depressed patients. Acta Psychiatrica Scandinavica, 96, 254-259.
- Vetter, P., & Koller, O. (1996). Clinical and psychosocial variables in different diagnostic groups: Their interrelationships and value as predictors of course and outcome during a 14-year follow-up. Psychopathology, 29, 159-168.
- VHA Directive 97-059 (1997). Instituting Global Assessment of Function (GAF) Scores in Axis V For Mental Health Patients. Washington D.C.: Department of Veterans Affairs.

- Walker, R., Minor-Schork, D., Bloch, R., & Esinhart, J. (1996). High risk factors for rehospitalization within six months. Psychiatric Quarterly, 67, 235-243.
- Walton, S. A., Berk, M., & Brook, S. (1996). Superiority of lithium over verapamil in mania: A randomized, controlled, single-blind trial. Journal of Clinical Psychiatry, 57, 543-546.
- Westermeyer, J., & Neider, J. (1988). Social networks and psychopathology among substance abusers. American Journal of Psychiatry, 145, 1265-1269.